



KÉPAF 2015



Képfeldolgozók és Alakfelismerők Társaságának
10. országos konferenciája

2015. Január 27-30. Kecskemét

**Képfeldolgozók és Alakfelismerők Társaságának
10. országos konferenciájának kiadványa**

Kecskemét 2015. január 27-30.

Tartalomjegyzék

Előszó	6
Bizottságok	7
KÉPAF Kuba Attila Díj	8
KÉPAF PhD Díj 2013-2014	9
Program	10
Elfogadott cikkek.....	16
Szóbeli szekció I.	16
Molnár József, Csetverikov Dmitrij, Kató Zoltán, Baráth Dániel: A Theory of Camera-Independent Correspondence	17
Tanács Attila, Majdik András, József Molnár, Atul Rai, Kató Zoltán: Megfeleltetés keresése sík képi régiók között	31
Hajder Levente, Baráth Dániel, Molnár József: Normálvektorok optimális becslése affin transzformációkból	44
Szóbeli szekció II.	60
Nagy Krisztián, Szirmay-Kalos László: Blind Logo and Identigram Exclusion for Image Matching	61
Csébfalvi Balázs: CC, BCC vagy FCC rács? 3D mintavételezés a Nyquist-küszöb alatt	71
Beleznai Csaba, Andreas Zweng, Peter Gemeiner: Reliable Left Luggage Detection Using Stereo Depth and Intensity Cues	80
Csordás Róbert, Havasi László, Szirányi Tamás: Kerítés fölött átdobott objektumok érzékelése kültéren	94
Technikai bemutatók (poszter szekció)	105
Báró Csaba, Alter Róbert: e_SENSOR valós idejű esemény detektáló videó analitikai rendszer bemutatása	106
Hajder Levente, Rui Huang, Kató Zoltán, Laczik Sándor, Majdik András, Molnár József, Molnár Zoltán, Rácz Péter, Atul Rai, Santa Zsolt, Tanács Attila: Collaborative 3D Reconstruction of Planar Surfaces from Mobile Cameras on Android Devices	110
Beleznai Csaba, Andreas Zweng: Stereo depth based human detection and tracking	114
Czúni László, Metwally Rashad, Kiss Péter, Gál Mónika, Lipovitz Ágnes: Lightweight Object Segmentation with Segmentation.....	117
Hegedűs Zoltán, Pintér István: Gépjármű motorszelep kimozdulás-idő függvényének mérése gyorskamerával	121

Szóbeli szekció III: Kuba Attila Díj.....	129
Koós Krisztián, Molnár József, Horváth Péter: DIC Microscopy Image Reconstruction Using a Novel Variational Framework.....	130
Tóth Márton, Csébfalvi Balázs: Agyszegmens detektálása CT felvételeken inverz anizotrop diffúzióval	140
Kovács György, Donkó Tamás, Matics Zsolt, Czakó Bernadett, Garamvölgyi Rita, Bajzik Gábor: On the CT based estimation of muscle weights in living rabbits	151
Szóbeli szekció IV: Kuba Attila Díj.....	167
Sánta Zsolt, Kató Zoltán: Ad-hoc mobil kamerahálózatok kalibrációja	168
Frohlich Róbert, Tamás Levente, Kató Zoltán: Omnidirekcionális kamerák közötti homográfia becslése pontmegfeleltetések nélkül	183
Börcs Attila, Nagy Balázs, Benedek Csaba: Valós idejű járműdetekció LIDAR pontfelhősorozatokon.....	199
Varga Domonkos, Havasi László, Szirányi Tamás: Human detection in real scenes.....	215
Poszter szekció I.	225
Kovács Levente: Geo-címkézett képhalmazok feldolgozása bővített nézetek kollaboratív előállításához	226
Shadaydeh Maha, Szirányi Tamás: A lokális hasonlósági mérték továbbfejlesztése távérzékelte képek változásainak becsléséhez	240
Manno-Kovács Andrea, Szirányi Tamás: Több irányú, alaksablonok nélküli épület detekció légi képeken	250
Mattyus Gellért, Friedrich Fraundorfer: Aerial image mosaicking with online calibration - A feasibility study	262
Zarándy Ákos: Cellular sensor-processor array based vision system	272
Csetverikov Dmitrij, Eichhardt Iván, Jankó Zsolt: A Brief Survey of Image-Based Depth Upsampling	279
Berke József, Kozma-Bognár Veronika: Turbulens áramlások vizsgálati lehetőségei légi felvételek geometriai és spektrális szerkezeti tulajdonságai alapján	295
Tóth János, Kovács László, Hartangi Balázs, Kiss Csaba, Mohácsi András, Orosz Zoltán, Hajdu András: An Online System for Benchmarking Image Processing Algorithms.....	305
Nagy Balázs, Benedek Csaba, Jankó Zsolt: Mozgó személyek követése és 4D vizualizációja Lidar-alapú járáselemzéssel	317
Balázs Péter, Kalmár György: Alacsony költségű, intelligens, mobil alapú otthoni felügyelő rendszer	333

Poszter szekció II.	347
Radim Krupička, Zoltán Szabó, Slávka Vítečková, Evžen Růžička: Motion Capture System for Finger Movement Measurement in Parkinson Disease	348
Fazekas Zoltán, Hajdu András, Lázár István, Kovács György, Csákány Béla, Dan Calugaru, Rajiv Shah, Ema Adam, Stefan Talu: Influence of Using Different Segmentation Methods on the Fractal Properties of the Identified Retinal Vascular Networks in Healthy Retinas and in Retinas with Vein Occlusion	360
Tóth Balázs, Papp László, Jakab Gábor, Patay Gergely, Szirmay-Kalos László: Adaptív bilaterális szűrés PET rekonstrukcióhoz	374
Szécsi László: GPU Pattern Generation for Retina Stimulation Experiments	388
Bognár Gergő: Képminőségmérés tüdő CT felvételeken	394
Horváth András, Dömötör Gábor: Egy retina-szimulációs szoftver és lehetőségek alkalmazásai	408
Szkalisity Ábel, Horváth Péter: A novel concept for high-content screen analysis using multi-parametric regression models to predict continuous cellular processes	418
Tamás Levente, Kató Zoltán: 3D Lidar-perspektív kamera alakzat alapú kalibrációja..	432
Kiss Attila, Szirányi Tamás: Szenzorhálózatok 3D összefűzése	444
Baráth Dániel: Homográfiabecslés affin transzformációból	458
Szóbeli szekció V.	474
Szirmay-Kalos László, Magdics Milán, Tóth Balázs: PET rekonstrukció javítása kívülről vezérelt anizotróp diffúzióval	475
Kevin Smith, Horváth Péter: Active Learning Strategies for Phenotypic Profiling of High Content Screens	489
Horváth András, Spindler Szabolcs, Rácz István, Szálai Milán: Kolonoszkópiás polipképek szoftveres kategorizálása	505
Szóbeli szekció VI.	515
Szirányi Tamás, Shadaydeh Maha : Távérzékelési képek szegmentálása MRF fúziós modellben hasonlósági mérték alapján	516
Benedek Csaba: Hierarchikus jelölt pontfolyamat modell objektumpopulációk többszintű elemzéséhez	528

Poszter szekció III.	538
Renner Gábor, Gyurecz György: Refinement of Shape and Structure of Highlight Lines	546
Bodnár Péter, Nyúl László, Grósz Tamás, Tóth László: Vizuális kódok lokalizációja mély egyenirányított neurális háló használatával	546
Katona Melinda, Nyúl László: Vonalkódok és természetes vizuális azonosítók felismerése valós időben.....	562
Németh Gábor, Palágyi Kálmán, Kardos Péter: Vékonyítás a végpont-megőrzés felülvizsgálatával	558
Kardos Péter, Palágyi Kálmán: Topology Preserving Reductions and Additions on the Triangular, Square, and Hexagonal Grids.....	588
Kiss Péter, Czúni László, Gál Mónika, Lipovitz Ágnes: Pivot-based search for word spotting in archive documents.....	601
Salvi Péter, Várady Tamás: Comparison of Two n-Patch Representations in Curve Network-Based Design	612
Kovács István, Salvi Péter, Várady Tamás: Strukturális alaksajátosságok érvényesítése CAD modellek rekonstrukciója során	625
Balázs Péter, Hantos Norbert: Polynomial Time Heuristics for Eliminating Switching Components via 0-1 Flips	641
Kovács György: On the Optimal Quantization Technique of Photometric Invariant Template Matching and Its Properties in the Presence of Gaussian Noise.....	657
Szóbeli szekció VII.	673
Vaitkus Márton, Várady Tamás: Labeled parameterization for high-quality surface fitting	374
Szécsi László, Szirányi Marcell: Dinamikusan generált textúra alapú vonalkázás	687
Szóbeli szekció VIII.	703
Manno-Kovács Andrea: Irányított vektormező konvolúció objektum körvonalak keresésére	704
Bodnár Péter, Nyúl László: QR kód lokalizáció kaszkádolt gyenge osztályozók használatával	712
Sávoli Zsolt, Horváth András, Kránicz Balázs: Színminta-halmazhoz illeszkedő, jó kvalitatív tulajdonságú metamer gyors keresése	722

Alacsony költségű, intelligens, mobil alapú otthoni felügyelő rendszer¹

Kalmár György¹, Balázs Péter²

¹ Szegedi Tudományegyetem, Természettudományi és Informatikai Kar
6720 Szeged, Aradi vértanúk tere 1.
Kalmar.Gyorgy@stud.u-szeged.hu

² Szegedi Tudományegyetem, Képfeldolgozás és Számítógépes Grafika Tanszék
6720 Szeged, Árpád tér 2.
pbalazs@inf.u-szeged.hu

Absztrakt. A betörések, rablások megelőzése vagy megakadályozása a hétköznapi ember számára csak riasztó berendezés használatával lehetséges, amely viszont drága és beszereléséhez szakemberre van szükség. Kutatásunk célja egy olyan alacsony költségű, mindenki számára elérhető, könnyen beüzemelhető és használható, mobil alapú védelmi berendezés megalkotása, amely képes megtanulni, majd később felismerni a védelmezett terület tulajdonosának arcát, felhasználva a ma már olcsón beszerezhető új, illetve használaton kívül esett régi okostelefonokat. A kifejlesztett felügyelő rendszer jelenleg egy arcot képes megtanulni, majd azt a későbbi megjelenés esetén felismerni. Ismeretlen arc esetén MMS vagy e-mail üzenetet küld a felhasználójának az arcról készült képpel. A szakirodalomban már jól kiforrott módszereket megfelelően egy működő rendszerre integráló Android alapú alkalmazás lehetővé teszi arcról nagy felbontású képek nagyon gyors kimentését, ezzel teret adva arcfelismerő algoritmusok használatának. A cikkben a kezdeti elért eredményeinket ismertetjük.

Kulcsszavak: megfigyelő rendszer, mobil alkalmazás, arcfelismerés

1 Bevezetés

Napjaink egyik kiemelkedő társadalmi problémája a betörések, lopások, rablások megakadályozása. Habár széles körben elérhetők otthonilag használható riasztóberendezések, ezek beszerelése sokszor túl költséges egy család számára. Továbbá egy rövid kódon múlik biztonságuk, amelyet akkor is igényelnek hangos riasztás közepette, ha a valódi tulajdonos érkezik otthonába. Megállapítható tehát, hogy ezek a rendszerek sokszor alacsony tudásúak és kényelmetlenek. Ezzel szemben egy okostelefon, amely ma már mindenki számára a hétköznapiak része, a

¹ Jelen kutatási eredmények megjelenését „Telemedicina fókuszú kutatások Orvosi, Matematikai és Informatikai tudományterületeken” című, TÁMOP-4.2.2.A-11/1/KONV-2012-0073 számú projekt támogatja. A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

kommunikáció és multimédia egész tárházát nyújtja egy tenyérnyi kis eszközben, magas processzor teljesítménnyel, sok memóriával, kamerával, mikrofonnal, internet hozzáféréssel, telekommunikációs funkciókkal.

Kutatásunk célja, egy olyan mindenki számára elérhető, olcsó, intelligens otthoni felügyelő rendszer kivitelezhetőségének a vizsgálata és megalkotása, amely képes megtanulni a rendszer használójának arcát, majd későbbi ismételt megjelenés esetén felismerni azt, ismeretlen személy esetén pedig riasztást küldeni MMS vagy e-mail üzenet formájában, amely tartalmaz egy képet az azonosítatlan arcról. Ezen rendszer elsősorban az elavulttá vált, sérült okostelefonok segítségével rendkívül kis költségigénnyel járulna hozzá a hétköznapi emberek biztonságához.

Ma már az okos eszközök nagyon nagy hányadán Android alapú operációs rendszer fut. Ez lehetővé teszi, hogy az egyes eszközök különbségeit a felsőbb rétegekben elrejtjük. Tehát egy elkészült, védelmi célokat biztosító alkalmazás az összes Android alapú eszközön futni képes. Ezen megfontolásból a megvalósításához mi is Android alapú eszközt választottunk és az elkészült alkalmazást is Java nyelven, Androidos környezetben fejlesztettük ki.

2 A megvalósított rendszerről általánosságban

A jelen megvalósításban tárgyalt rendszer csak egy arc megtanulására képes. Ezen korlátozás kihatással van az alkalmazott módszerre. A mesterséges intelligencia megalkotásánál figyelembe kell venni, hogy a döntéshozás bináris jellegű, tehát megtanult arc vagy nem az szerepel egy képen.

Az arcfelismerés statikus képek esetén bár jól, de nem tökéletesen megoldott probléma. Esetünkben a felismerés nehézségéhez hozzájárul az is, hogy nem élhetünk azzal a feltételezéssel, hogy a felvétel az arcról szemből készül. Ezért már egy személy felismerésénél is adódnak nehézségek. Elsősorban az intraperszonális variancia, vagyis hogy egy személy saját magához képest mennyire változatos, jelenti a fő nehézséget. Akadályozó tényező továbbá, hogy esetünkben a felismerendő személyek mozognak, így csak homályos képeket készíthetünk, ha figyelembe vesszük, hogy az okostelefonok kamerája nem a legkifinomultabb optikai rendszer és nem is a gyors képkészítésre optimalizált.

A képfeldolgozási algoritmusokat az alkalmazás az OpenCV4Android nevű csomagon keresztül éri el, amely az OpenCV Androidos környezetre előkészített változata. Az eljárások nagy része C++-ból fordított natív kód formájában kerül futtatásra a Java-s környezetből egy JNI (Java Native Interface) csatoláson keresztül, ezzel lényegesen felgyorsítva a végrehajtást. Ezen rendszerben rendelkezésre álló lehetőségekről részletes tájékoztatást nyújt az [1], amelyből azonban sok eljárás nem érhető el az Androidos verzióban.

A rendszerrel szemben elvárt funkciók:

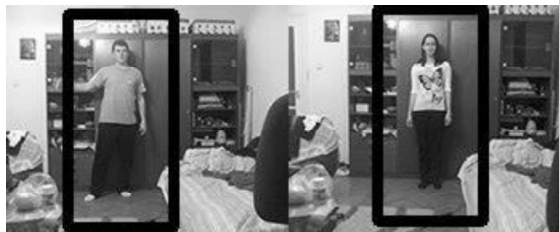
- legyen *képes* egy olcsó okostelefonon is futni (minimális elvárás, hogy az rendelkezzen egy hátoldali kamerával, egy mikrofonnal, és képes legyen Android 4.0-s vagy újabb verziót futtatni)

- legyen képes egy arc megtanulására, ezt egy egy napos, felügyelt tanítási fázisban kell, hogy megtegye
- a tanítási fázis után képes legyen egy megjelenő arcról eldönteni, hogy az a megtanult arc (személy) vagy sem
- mindezen felismerő funkciókat a lehető leghatékonyabban és legmegbízhatóbban végezze
- minimalizálja a fals pozitív döntések számát, hiszen ez azt jelentené, hogy ismeretlen arcot ismertnek feltételezne
- ha az arc felismerése fizikai korlátok miatt nem is hajtható végre, akkor is próbáljon meg döntést hozni az illető kilétéről, ezt jelen megvalósításban egy hangalapú jelszóellenőrző algoritmus végzi el

A rendszer betüzemelésének kezdetekor az adott okostelefonra fel kell telepíteni az alkalmazást. Elindítva a programot, az egyből a tanulási fázisba lép. Ezen fázis során tanulja meg a később felismerni kívánt arcot, illetve a hangalapú jelszó azonosításhoz a referencia jelszót. A tanulás végeztével a rendszer egyből riasztó, felügyelő üzembe lép. Ez az üzemmód az alkalmazás leállításáig tart.

3 A rendszer alapkövei

Az emberi alak megbízható és gyors felismerése az egyik, és talán a legfontosabb alappillére a felügyelő rendszernek. Ennek köszönhető, hogy a képen megjelenő személy detektálásra kerül, így a rendszer értesül arról, ha a megfigyelt térségben valaki tartózkodik. Jelenleg erre a célra, az OpenCV által is támogatott HOG leíróval működő módszert használjuk [2,3]. Ennek előnye, hogy nagyon megbízható, a fals pozitív értesítések száma minimális. Ugyanakkor igen műveletigényes, gyengébb processzoron és nagyobb felbontású képen futtatva több másodpercig is eltart az elemzés. Ennek kiküszöbölésére hoztuk létre a később tárgyalásra kerülő saját kamerakezelő osztályt, amely segítségével már a hardvertől érkező képek felbontását megfelelően alacsonyra tudjuk állítani ahhoz, hogy a HOG algoritmus szinte valós időben működhessen. Az emberi alak detektálásának működése során készült képek az 1. ábrán láthatóak.



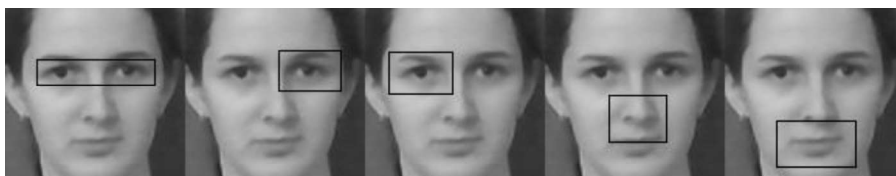
1. ábra: Emberi alak felismerés

Az emberi alak megtalálásán túl a rendszer másik elengedhetetlen eleme az arcok megbízható és gyors detektálása, tekintve, hogy a végső cél ezek megtanulása és

felismerése. Még az emberi alak többnyire egy általános formának mondható, addig az arc nagyon változatos kinézetű is lehet. Az OpenCV az eddig publikált leggyorsabb és legrobustusabb arcfelismerő eljárást támogatja. Ennek lényege, hogy úgynevezett Haar jellemzőket nyernek ki a képekből, majd ezekre építve gyenge osztályozókat hoznak létre, amelyeket egymás után fűznek (*Haar cascade classifier*) [4]. Az OpenCV ezen eljárásra támaszkodva, előre tanított osztályozókat biztosít, amelyek nem csak arc, de sok más objektum keresésére is alkalmasak [5]. Ennek köszönhetően választottuk az arcfelismerés módszeréül azt, hogy az arcról ún. mikrojellemzőket vonunk ki és ezeket hasonlítjuk össze a később megjelenő arc jellemzőivel. A mikrojellemzők konkrétan a száj, az orr, a bal szem, a jobb szem, és a szempár, pontosabban az ezekről készült képrészletek. Az arc és arcrészek detektálására példákat a 2. és 3. ábrán láthatunk.



2. ábra: Arc detektálása



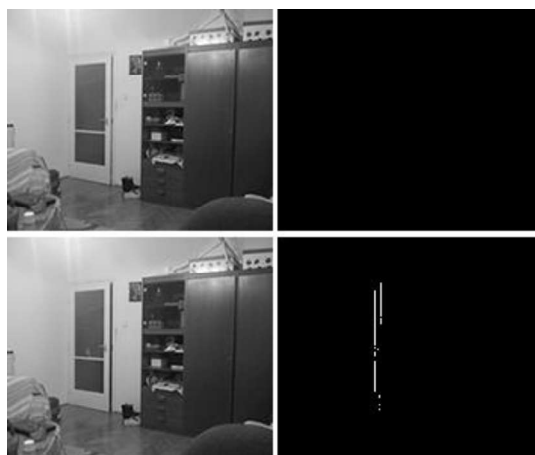
3. ábra: Arcrészek detektálása

4 Kiegészítő funkciók

A felügyelő rendszer alapvető funkcióinak tárgyalása után a következőkben bemutatjuk a nem szükségszerű, de nagyon hasznos kiegészítő eljárások működését. Ezek közé tartozik a pehelysúlyú mozgásérzékelő, a hangalapú jelszófeldolgozó, a dinamikus arckövető algoritmus és a saját kamerakezelő osztály.

Habár az emberi alak keresés nagyon hatékony a kis méretű képen, és így a feldolgozási sebességet is magasan lehet tartani, mégis indokolatlan egy ilyen bonyolultságú algoritmust futtatni minden képre. Ha ugyanis nincs mozgás a képen, akkor biztosak lehetünk benne, hogy ember nem sétált be a megfigyelt területre. Ezt

kihasználva egy mozgásérzékelő algoritmus előszűrője lehet az emberi alak keresésnek, miszerint azt csak akkor érdemes futtatni, ha történt a képen jelentős változás. Nyilvánvaló, hogy ez az algoritmus sokkal egyszerűbb, kisebb processzor és memóriaigénnyel. A szakirodalom *háttérkivonás* néven hivatkozik az olyan algoritmusokra, amelyek célja, hogy egy képen elkülönítsék a háttérhez és az objektumhoz (előtérhez) tartozó pixeleket. A legkiemelkedőbb módszerek ezen a téren valószínűségi háttér modelleket alakítanak ki. Az általunk használt, OpenCV-ben megvalósított módszer lényege, hogy minden pixel intenzitását egy tanító képhalmaz segítségével egy normális eloszlással adja meg. Vagyis minden pixelre meghatároz egy várható értéket és szórást, és az eloszlást normálisnak feltételezi (változó paraméterű Gauss görbékkel ír le minden pixelt). Innen ered a neve, *Gaussian Mixture Model*, vagyis Gauss görbék keverékéből álló modell [6]. Az OpenCV-ben szereplő verzió nagyon gyors és műveletigénye is alacsony [7]. Működés közben készült képeket a 4. ábrán figyelhetünk meg.



4. ábra: Mozgásérzékelés

Az eddig tárgyalt funkciók célja, hogy lehetővé tegyék a felügyelő rendszer számára, hogy érzékelje az emberi jelenlétet és megpróbálhassa a személy arcát felismerni. Számos okból előfordulhat azonban, hogy az arc detektálásának fizikai okokból eleve lehetetlen (pl.: rossz fényviszonyok, eltakart vagy a kamerának háttal forduló arc, gyors mozgás, amely miatt lehetetlen éles képet készíteni). Ilyen esetekben is valahogy meg kellene arról bizonyosodni, hogy a detektált emberi alak ismert archoz tartozik-e.

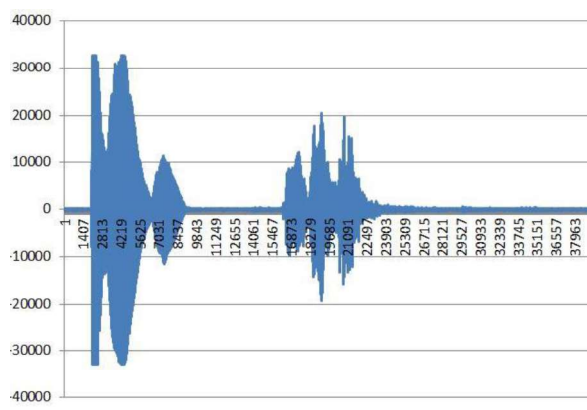
A probléma megoldásához a hang alapú interakciót választottuk. Az azonosítás egy előre megtanult jelszó ismételt kimondásával valósul meg. A tanulási fázis elején a felhasználó egy tetszőleges jelszót mondhat be a rendszerbe, amit az felvesz és megjegyez. Ha a későbbi éles működés során az arc felismerése sikertelen lenne valamilyen okból kifolyólag, akkor a rendszer egy hangjelzés segítségével „megkéri” a felügyelt területen tartózkodó személyt, hogy ismétlje el a megtanult jelszót. Ezt felvételezi, majd kinyeri a felvételi idő beszéd részét. A referencia és az aktuálisan

meghatározott beszéd régiók korrelációja alapján döntést hoz, hogy ugyanaz a szó hangzott-e el.

Az eszköz által felvett jel nagyon zajos lehet. Erre láthatunk példát az 5. ábrán. A jelen egyenirányítást majd szűrést hajtunk végre a következő szűrő segítségével (csúszóablakos, súlyozott, átlagoló szűrő):

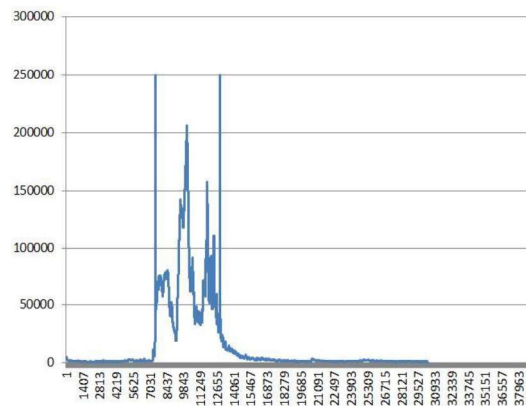
$$f(x) = \frac{1}{2n+1} \sum_{i=-n}^n (n - |i| + 1) * y(x+i), x = n, \dots, L-n$$

ahol $f(x)$ a szűrt jel az x pozícióban, $n=N/2$, ahol N a csúszóablak szélessége. L az eredeti jel hossza, $y(x)$ pedig az értéke az x pozícióban.



5. ábra: Nyers hangfelvétel hullámformája

Az 5. ábrán látható jelalak elején megfigyelhető kiemelkedések a rendszer által kiadott hang miatt jelennek meg, ami mindig ugyanolyan hosszú, tehát eldobható ez a rész. Az 5. ábra jelének szűrt, kezdeti részt nem tartalmazó alakja a 6. ábrán figyelhető meg, amelyen a két kiemelkedő csúcs az algoritmus által meghatározott beszéd régió kezdő és végpontját jelöli. A beszéd régió jelen megvalósításban a leghosszabb, nem zajszinten lévő régió.



6. ábra: Szűrt jel, bejelölt beszéd régióval

Ha már rendelkezésre állnak a beszéd régiók, akkor ezeket kinyerve a referencia és az aktuálisan felvett hangból, kezdődhet az összehasonlítás. Ennek során az összehasonlítandó beszédrészt ugyanolyan hosszúra alakítjuk, mint a referencia jelszó hossza (nyújtás, majd interpoláció). Ez azért fontos, mert ugyanazt a szót többféle ritmusban is ki lehet mondani. Ha már egyforma hosszúak, normalizált keresztkorrelációval kiszámítjuk az egyezést a két hangalapú jelszó közt. Ha egy előre megadott küszöböt meghalad a hasonlóság mértéke, akkor a bediktált jelszót elfogadjuk, egyébként pedig elvetjük.

Már az előzőekben említés esett arról, hogy fontos és elengedhetetlen az arcról nagyon gyorsan, a lehető legnagyobb felbontással képeket készíteni. Ehhez szükségünk van egy olyan eljárásra, amely biztosítja, hogy ne a teljes képet kelljen átkutatni arc után. Ehhez két feltevéssel élhetünk:

1. Emberi alak fellelése után elegendő csak az emberi alak helyének felső részén keresni arc után egy adott méretű régióban.
2. Ha már találtunk arcot, valószínű, hogy a következő képen is ehhez a helyhez nagyon közel lesz az arc.

Ezeket kihasználva egy olyan módszert fejlesztettünk ki, amely nyomon tudja követni az arc helyzetét. Az eljárás egy növekvő méretű keresési ablakot használ. Az ablak mérete azzal arányosan nő, hogy hány képkocka óta nem volt arc az adott ablakban. Ha ismét talál egy arcot, akkor az ablak középpontját áthelyezi az arc középpontjába és innen folytatja újra a növekedést, ha ismét eltűnne az arc. A 7. ábrán egy olyan sorozatot láthatunk, amelynek első képén az ablak az emberi alak detektálás után állítódott be. Majd a további képeken a mozgás hatására változó helyzetű és méretű ablak figyelhető meg.



7. ábra: Dinamikus arckövetés

Az alkalmazás szempontjából kritikussá vált, hogy elérhetőek legyenek a következő, kamerát érintő funkciók:

1. Változtatni lehessen a kamerából érkező kép felbontását.
2. Folyamatosan bekapcsolva lehessen tartani a felvétel során a kamerához tartozó LED vakut (éjjeli üzemmód).

Ezen funkciókat csak saját kamerakezelő osztály megírásával lehet elérni. Ennek megvalósítását követően a felbontás váltása hardver szinten gyorsan, kevesebb, mint egy másodperc alatt végbemegy, illetve folyamatos üzemben tudjuk működtetni a kamerához tartozó LED vakut.

5 A rendszer működési fázisainak vizsgálata

A rendszer feladata, hogy megjelenő arcokat „ismert” vagy „ismeretlen” címkével lásson el. Ehhez azonban először egy felügyelt tanulási fázison kell átesnie, amely során elsajátíthatja a később felismerendő arcot. Ezt az alkalmazás első indításakor teszi meg. Maga a tanulás a később felismerni kívánt jellemzők kivonását, elmentését jelenti. Jelen megvalósításban ezek a leírók az arcról kivont mikrojellemzők: száj, orr, bal szem, jobb szem, szempár. Ez az öt kép kerül elmentésre egy referencia arcról, minden arcrész egy előre megadott egységes méretre skálázva. A rendszer ezen mikroképeket egy napig gyűjti, fontos, hogy ez az idő alatt csak a később felismerni kívánt felhasználó jelenjen meg a felügyelt területen.

A képkészítések során homályos, elmosódott képek sokszor készülnek. Ennek kiszűrése érdekében egy referencia helyett több kerül elmentésre minden arcrészről. A korrelációt zavarja a fény intenzitásának és irányának változása, ezért ezt kiküszöbölendő, a rendszer egy nap során, több alkalommal is készít referenciákat, időben elszeparálva. Ha egy lokális referencia gyűjtési időben már megvan a kellő számú minta minden mikrojellemzőről, akkor a rendszer tovább folytatja a képek gyűjtését, és az itt kivont jellemzőket azonnal össze is hasonlítja az éppen rögzített referencia jellemzőkkel. Egy személy adott arcrészét sajátjával hasonlítjuk össze.

Ezen autokorrelációs értékekből sok készül el. Ezekből átlagot számolva egy dinamikusan meghatározott küszöbszintet állítunk elő a későbbi felismerés számára, melyet időben hozzárendelünk a referenciákhoz. Egy referenciacsoportot és a hozzá rendelt küszöbszinteket láthatjuk a 8. ábrán.



8. ábra: Rögzített referenciacsoport dinamikusan meghatározott küszöbszintekkel.

Amikor a tanulás véget ér, a felügyelő fázis kezdi meg munkáját. Ebben a módban, a rendszer folyamatosan, bizonyos időnként referencia betöltést hajt végre. Ez a dinamikus referencia betöltés. Ennek során az elmentett referenciákat végignévezve dönt arról, hogy a betöltés idején (valószínűsíthetőleg) melyek azok, amelyek a legjobban illenek az akkori fényviszonyokhoz (intenzitás, irány).

A felügyelés során folyamatos mozgásérzékelés történik. Ha mozgás jelenik meg a képen, akkor a rendszer egyből emberi alak detektálásra vált át. Ha nem talál embert, akkor bizonyos idő után visszaáll mozgásérzékelésre. Ha talál, akkor a dinamikus arckövető eljárással arc után kezd kutatni. Ha nincs arc, akkor jelszót kér bemondásra, hiszen arc nélkül a felismerő rendszer nem működhet.

Arc detektálása után az arról készült kép elmentése kerül. Adott számú mentett arc után a döntéshozás algoritmusát lát munkához. A referencia és aktuálisan kivont mikroképek normalizált keresztkorrelációja kiszámításra kerül a következő formula alapján:

$$R(x, y) = \frac{\sum_{x', y'} (T(x', y') * I(x + x', y + y'))}{\sqrt{\sum_{x', y'} T(x', y')^2 + \sum_{x', y'} I(x + x', y + y')^2}}$$

ahol T a referencia kép, I pedig az aktuális, összehasonlítandó kép. Esetünkben egy korrelációs értéket kapunk az összehasonlítás eredményeként, mivel a képeink megegyező méretűek. Innen egy lokális döntés:

$$\text{lokális döntés} \begin{cases} \text{igen} & \text{ha } R(T, I) \geq \text{küszöbszint} \\ \text{nem} & \text{ha } R(T, I) < \text{küszöbszint} \\ \text{ismeretlen} & \text{ha nem sikerült kinyerni a jellemzőt} \end{cases}$$

ahol $R(T, I)$ a referencia és az összehasonlítandó kép keresztkorrelációjának értéke, a *küszöbszint* pedig a tanulási fázisból dinamikusan, autokorrelációkból származtatott érték. Innen egy arcra vonatkoztatott döntés:

$$\text{arc szintű döntés} \begin{cases} \text{extra igen} & \text{ha } i \geq e * n \\ \text{igen} & \text{ha } i > t * n \\ \text{nem} & \text{ha } i < t * n \\ \text{extra nem} & \text{ha } i < (1 - e) * n \\ \text{ismeretlen} & \text{ha } u \geq (1 - e) * n \end{cases}$$

ahol i azon lokális döntések száma, amelyek eredménye „igen”, e egy magas érték a $[0,1]$ -ből, t alacsony érték $[0,1]$ -ből, $n = r_f * n_f$, n_f a jellemzők száma, r_f a jellemzőnként rögzített képek száma, u az „ismeretlen” lokális döntések száma. Jelenlegi megvalósításban: $n_f=5$, $r_f=5$, $e=0,8$, $t=0,5$. Innen a végleges döntés:

$$\text{végleges döntés} \begin{cases} \text{igen} & \text{ha } e_i \geq w * f \\ \text{nem} & \text{ha } e_n \geq w * f \\ \text{ismeretlen} & \text{ha } e_i = 1 \text{ és } n_i < t * (f - 1) \\ \text{igen} & \text{ha } e_i = 1 \text{ és } n_i \geq t * (f - 1) \\ \text{ismeretlen} & \text{ha } e_n = 1 \text{ és } n_n < t * (f - 1) \\ \text{nem} & \text{ha } e_n = 1 \text{ és } n_n \geq t * (f - 1) \\ \text{igen} & \text{ha } n_i \geq h * f \\ \text{nem} & \text{ha } n_n > (1 - h) * f \\ \text{ismeretlen} & \text{egyébként} \end{cases}$$

ahol: f az arci képek (döntések) száma, e_i az „extra igen”, e_n az „extra nem”, n_i az „igen”, n_n a „nem” arconkénti döntések száma, w egy $[0,1]$ -ből származó kis értékű szám, t ugyanaz, mint korábban, h egy $[0,1]$ -ből származó szám. Jelenlegi megvalósításban $h=t$.

A döntéshozó „nem” kimenete esetén a rendszer egy e-mail illetve MMS üzenet küldését vonja maga után, amely tartalmaz egy képet az ismeretlen arcról. A riasztást követően ismét mozgásérzékeléssel folytatódik a felügyelet.

A döntéshozó „ismeretlen” kimenete esetén a rendszer hangjelzéssel kéri a megfigyelt területen tartózkodó személyt, hogy mondja be az azonosításhoz szükséges jelszót. Ennek ellenőrzésével teszi „igenné” vagy „nemmé” a bizonytalan kimenetet. Ha a jelszó nem megfelelő, a fentebb leírt riasztás indul el.

A döntéshozó „igen” kimenetére a rendszer nyugtázza a döntést és mozgásérzékeléssel folytatja munkáját. Egy elkészült döntést tartalmazó kép a 9. ábrán tekinthető meg (a képen a bal felső sarokban az éppen aktuális küszöbszintek arcrészek szerint feltüntetve, alatta a végleges döntés, alatta a lokális döntések, arconként oszlopba gyűjtve láthatók).



9. ábra: Döntéshozás eredménye

6 A rendszer tesztelése

Az elkészült felügyelő rendszer tesztelése több módon és fázisban zajlott. Külön-külön teszteltük az egyes komponensek megbízhatóságát, majd az együttműködésük eredményét. Némely fázist több készüléken is vizsgáltunk. Ezek közt szerepelt: Pulid F11, 2x1GHz elérhető maximális felbontás 1280x720, ár: 20 000Ft (kínai); LG G3, 4x2,5GHz, használt felbontás 2048x1080, ár: 150 000Ft; Samsung S Advance, 2x1GHz, elérhető maximális felbontás 800x600, ár: 40 000Ft,

Az egyes komponensek vizsgálata során mindegyik gyors működést és nagyon jó eredményeket ért el. Magának az arcfelismerő algoritmusnak a tesztelése volt a másik fő feladat. Az ütemezett referencia készítésnek és betöltésnek a tesztelése sikeres volt. A személyek felismerését több módon is teszteltük. Az első tesztelési forma statikus állapotban, a referencia készítésének helyén történő mérés volt. A statikus állapot a személyre igaz, vagyis nem mozgott a képek készítése során, és a referencia készítés helyén és időpontjában történt a tesztelés. Ez a legjobb eshetőség, hiszen pontosan azokat az anomáliákat szünteti meg, amelyek a korrelációt megzavarják. A tesztek során a felhasználó szerepelt a képeken, tehát a válasz mindig „igen” kellett volna, hogy legyen. Az eredmények az 1-4. táblázatokban figyelhetők meg.

Eszköz	Mérések száma	Igen	Nem	Talán (megerősítve hang alapon)
Pulid F11	30	18	4	8 (6)
Samsung S Advance	20	10	3	7 (6)
LG G3	15	8	4	3 (0)

1. táblázat: Statikus, megtanult személy felismerése

Eszköz	Mérések száma	Igen	Nem	Talán (megerősítve hang alapon)
Pulid F11	20	3	10	7 (1)
Samsung S Advance	20	5	11	4 (0)
LG G3	0	0	0	0

2. táblázat: Statikus, ismeretlen személy felismerése.

Eszköz	Mérések száma	Igen	Nem	Talán (megerősítve hang alapon)
Pulid F11	20	5	6	9 (7)
Samsung S Advance	20	7	7	6 (6)
LG G3	10	2	4	4 (3)

3. táblázat: Dinamikus, ismert személy felismerése.

Eszköz	Mérések száma	Igen	Nem	Talán (megerősítve hang alapon)
Pulid F11	20	1	3	16 (2)
Samsung S Advance	20	2	4	14 (0)
LG G3	10	2	3	5 (1)

4. táblázat: Dinamikus, nem ismert személy felismerése

A nehezebb felismerési forma, amikor a személy nem vesz tudomást a készülékről és nem is néz bele, illetve mozgást végez, amely miatt a képek sem élesek. Nyilvánvaló, hogy itt nagyobb szerepet kap a hang alapú ellenőrzés. Maga a tesztelés egy megtanult mozgási sor végrehajtásával történt, hogy az eredmények összehasonlíthatóak legyenek. A 3. táblázatban szereplő mérések esetén a rendszer felhasználója, a megtanult személy felismerése volt a cél. A 4. táblázatban a mozgó, ismeretlen személy alapú mérések összegzése található.

A tesztelés eredményeit röviden összefoglalva elmondható, hogy az eszköz gyorsasága és minősége nem hat ki lényegesen az eredményekre. Az egyértelmű, hogy statikus helyzetben a rendszer megfelelően működik. Az idegen személyek azonosítása kicsivel gyengébb eredményeket hozott, de ha az illető nem ismeri a jelszót, akkor megfelelő a működés, hiszen a hang alapú azonosítás ilyen esetben a főszerep (de pontosan ezért is lett implementálva). Az eredményeket tekintve az elvárt működést tapasztaltuk, az ismert hiányosságok és korlátok jól látszódnak.

Kiforrott módszerek [8,9,10,11] pontosságának elérése fejlettebb eljárások bevonását igényli a jövőben, illetve ezek alkalmazását is érdemes fontolóra venni.

7 Összefoglalás

Kutatásunk célja egy olyan mobil alapú, intelligens, alacsony költségű felügyelő rendszer megalkothatóságának a vizsgálata és létrehozása volt, amely képes felismerni a felhasználója arcát. Sikertült egy olyan rendszert megalkotnunk, amely lehetővé teszi arcokról készült képek gyors gyűjtését, ezzel teret adva további, mélyebb kutatásoknak. A jelen megvalósításban szereplő felügyelő rendszer az arcról kivont mikrojellemzők (száj, orr, bal szem, jobb szem, szempár) képtérbeli összehasonlításával ítéli meg, hogy ismert arc szerepel-e egy adott képen. Habár sikerült így is megfelelő eredményeket elérnünk, nyilvánvaló, hogy ezek a jellemzők nagyon érzékenyek a varianciákra. További fejlesztésként elsősorban ennek javítását tűzzük ki célul. A megvalósított algoritmusok közül különösen jól működött a hang alapú jelszóellenőrzés, amely érzékenységét a külső zajokra még javítani kell, de így is 85% feletti pontosságot ért el a tesztelés során.

A rendszer egyszerűen bővíthető, újabb jellemzők vizsgálata, bevétele könnyedén megoldható. A legtöbb eljárás paraméterezhető, tehát további kutatás folyamán is segítséget nyújt a szabad kísérletezésre. Az alkalmazott detektáló eljárások nagyon hatékonyan működnek, ezzel lehetővé téve az arcok gyors rögzítését. A saját kamerakezelő osztály további funkciók kifejlesztését teszi lehetővé, mint például az éjjeli üzemmód, a LED vaku ugyanis szabadon kapcsolható programból. Az alkalmazás Android alapokon íródott, így szinte tetszőleges eszközre változtatás nélkül feltelepíthető és azonnal használható. Az arcfelismerő funkció jelenleg statikus helyzetben lévő személyek esetén megfelelően biztonságosan, még mozgó egyének esetén, érthető korlátok miatt gyengébben szerepel. A tapasztalatok alapján úgy gondoljuk, hogy lehetséges egy olyan rendszer megalkotása, amely közel minden esetben képes egy adott arcról megfelelő döntést hozni. További bővítési és fejlesztési lehetőségek bőven adódnak: pontosított arcfelismerés, több megtanult arcra működő azonosítás, gépi tanulással támogatott döntéshozás, fejlettebb interakciós formák a felhasználóval, hogy csak néhányat említsünk. A rendszer tömegesebb alkalmazása esetén segítené felhasználni az egyébként kihasználatlan okostelefonokat és egyben segítené megóvni otthonainkat, értékeinket és támogatná a bűnüldözést.

Irodalomjegyzék

1. Daniel Lélis Baggio, Shervin Emami, David Millán Escrivá, Khvedchenia Ievgen, Naureen Mahmood, Jason Saragih, Roy Shilkrot: Mastering OpenCV with Practical Computer Vision Projects, Birmingham, UK. Packt Publishing Ltd., November 2012

2. N. Dalal, B. Triggs, Histogram of oriented gradients for human detection, IEEE Conference on Computer Vision and Pattern Recognition, pp. 886-893, 2005.
3. http://docs.opencv.org/modules/gpu/doc/object_detection.html
4. P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features. IEEE Conference on Computer Vision and Pattern Recognition, pp. 511-518, 2001.
5. http://docs.opencv.org/modules/objdetect/doc/cascade_classification.html
6. Z. Zivkovic, Improved adaptive Gaussian mixture model for background subtraction, IEEE International Conference on Pattern Recognition, pp. 28-31, 2004.
7. <http://docs.opencv.org/java/org/opencv/video/BackgroundSubtractorMOG2.html>
8. R. Brunelli, T. Poggio, Face recognition: Features versus templates. IEEE Trans. Pattern Analysis and Machine Intelligence 15(10), pp. 1042-1052, 1993.
9. L.-F. Chen, H.-Y.M. Liao, M.-T. Ko, J.-C. Lin, G.-J. Yu, A new LDA-based face recognition system which can solve the small sample size problem. Pattern Recognition 33(10), pp. 1713-1726, 2000.
10. A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, Y. Ma, Toward a practical face recognition system: Robust alignment and illumination by sparse representation. IEEE Trans. Pattern Analysis and Machine Intelligence 34(2), pp. 372-386, 2012.
11. F. Goudail, E. Lange, T. Iwamoto, K. Kyuma, N. Otsu, Face recognition system using local autocorrelations and multiscale integration. IEEE Trans. Pattern Analysis and Machine Intelligence 18(10), pp. 1024-1028, 1996.